

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

(i) APPLICANTS: Morrow, Casey D. and Porter, Donna, C.

10 (ii) TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS
NUCLEIC ACID AND METHODS OF MAKING AND
USING SAME

(iii) NUMBER OF SEQUENCES: 23

15 (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 60 STATE STREET, SUITE 510
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS
20 (E) COUNTRY: USA
(F) ZIP: 02109

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
25 (B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US 000000
30 (B) FILING DATE: 15-FEB-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 08/087,009
35 (B) FILING DATE: 01-JUL-1993
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Silveri, Jean M.
40 (B) REGISTRATION NUMBER: P-39,030
(C) REFERENCE/DOCKET NUMBER: UAG-004CP

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (617) 227-7400
45 (B) TELEFAX: (617) 227-5941

50 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
55 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 TATTAGTAGA TCTG

14

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

20 TACAGATGTA CTAA

14

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 846 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- 35 (A) NAME/KEY: CDS
(B) LOCATION: 20..845

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

40 ACACAGCAAT CAGGTCAGC CAA AAT TAC CCT ATA GTG CAG AAC ATC CAG GGG 52
Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly
1 5 10

45 CAA ATG GTA CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA TGG GTA 100
Gln Met Val His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val
15 20 25

50 AAA GTA GTA GAA GAG AAG GCT TTC AGC CCA GAA GTG ATA CCC ATG TTT 148
Lys Val Val Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe
30 35 40

55 TCA GCA TTA TCA GAA GGA GCC ACC CCA CAA GAT TTA AAC ACC ATG CTA 196
Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu
45 50 55

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 275 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln
1 5 10 15
Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu
20 25 30
Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu
35 40 45
Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly
50 55 60
His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala
65 70 75 80
Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro
85 90 95
Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser
100 105 110
Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro
115 120 125
Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile
130 135 140
Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro
145 150 155 160
Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg
165 170 175
Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu
180 185 190
Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu
195 200 205
Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val
210 215 220
Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln
225 230 235 240
Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg Asn
245 250 255

Gln Arg Lys Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Thr
 260 265 270

5 Ala Arg Lys
 275

(2) INFORMATION FOR SEQ ID NO:5:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 948 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

20 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 4..946

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

25 AAC CAA TGG CCA TTG ACA GAA GAA AAA ATA AAA GCA TTA GTA GAA ATT 48
 Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 1 5 10 15

30 TGT ACA GAG ATG GAA AAG GAA GGG AAA ATT TCA AAA ATT GGG CCT GAA 96
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 20 25 30

35 AAT CCA TAC AAT ACT CCA GTA TTT GCC ATA AAG AAA AAA GAC AGT ACT 144
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 35 40 45

40 AAA TGG AGA AAA TTA GTA GAT TTC AGA GAA CTT AAT AAG AGA ACT CAA 192
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 50 55 60

45 GAC TTC TGG GAA GTT CAA TTA GGA ATA CCA CAT CCC GCA GGG TTA AAA 240
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 65 70 75

50 AAG AAA AAA TCA GTA ACA GTA CTG GAT GTG GGT GAT GCA TAT TTT TCA 288
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 80 85 90 95

55 GTT CCC TTA GAT GAA GAC TTC AGG AAG TAT ACT GCA TTT ACC ATA CCT 336
 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 100 105 110

AGT ATA AAC AAT GAG ACA CCA GGG ATT AGA TAT CAG TAC AAT GTG CTT 384
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 115 120 125

	CCA CAG GGA TGG AAA GGA TCA CCA GCA ATA TTC CAA AGT AGC ATG ACA	432
	Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr	
	130 135 140	
5	AAA ATC TTA GAG CCT TTT AGA AAA CAA AAT CCA GAC ATA GTT ATC TAT	480
	Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr	
	145 150 155	
10	CAA TAC ATG GAT GAT TTG TAT GTA GGA TCT GAC TTA GAA ATA GGG CAG	528
	Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln	
	160 165 170 175	
15	CAT AGA ACA AAA ATA GAG GAG CTG AGA CAA CAT CTG TTG AGG TGG GGA	576
	His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly	
	180 185 190	
20	CTT ACC ACA CCA GAC AAA AAA CAT CAG AAA GAA CCT CCA TTC CTT TGG	624
	Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp	
	195 200 205	
25	ATG GGT TAT GAA CTC CAT CCT GAT AAA TGG ACA GTA CAG CCT ATA GTG	672
	Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val	
	210 215 220	
30	CTG CCA GAA AAA GAC AGC TGG ACT GTC AAT GAC ATA CAG AAG TTA GTG	720
	Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val	
	225 230 235	
35	GGG AAA TTG AAT TGG GCA AGT CAG ATT TAC CCA GGG ATT AAA GTA AGG	768
	Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg	
	240 245 250 255	
40	CAA TTA TGT AAA CTC CTT AGA GGA ACC AAA GCA CTA ACA GAA GTA ATA	816
	Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile	
	260 265 270	
45	CCA CTA ACA GAA GAA GCA GAG CTA GAA CTG GCA GAA AAC AGA GAG ATT	864
	Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile	
	275 280 285	
50	CTA AAA GAA CCA GTA CAT GGA GTG TAT TAT GAC CCA TCA AAA GAC TTA	912
	Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu	
	290 295 300	
55	ATA GCA GAA ATA CAG AAG CAG GGG CAA GGC CTCGAG	948
	Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly	
	305 310	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 314 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5 Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys
1 5 10 15

Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn
20 25 30

10 Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys
35 40 45

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp
50 55 60

15 Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys
65 70 75 80

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val
85 90 95

Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser
100 105 110

25 Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro
115 120 125

Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
130 135 140

30 Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln
145 150 155 160

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
165 170 175

35 Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu
180 185 190

40 Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
195 200 205

Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu
210 215 220

45 Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
225 230 235 240

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln
245 250 255

50 Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro
260 265 270

55 Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
275 280 285

Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
 290 295 300

5 Ala Glu Ile Gln Lys Gln Gly Gln Gly Leu
 305 310

(2) INFORMATION FOR SEQ ID NO:7:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1568 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

20 (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 7..1565

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

25	GGGGCC TGT CCA AAG GTA TCC TTT GAG CCA ATT CCC ATA CAT TAT TGT Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys 1 5 10	48
30	GCC CCG GCT GGT TTT GCG ATT CTA AAA TGT AAT AAT AAG ACG TTC AAT Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn 15 20 25 30	96
35	GGA ACA GGA CCA TGT ACA AAT GTC AGC ACA GTA CAA TGT ACA CAT GGA Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly 35 40 45	144
40	ATT AGG CCA GTA GTA TCA ACT CAA CTG CTG TTA AAT GGC AGT CTA GCA Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala 50 55 60	192
45	GAA GAA GAG GTA GTA ATT AGA TCT GTC AAT TTC ACG GAC AAT GCT AAA Glu Glu Glu Val Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys 65 70 75	240
50	ACC ATA ATA GTA CAG CTG AAC ACA TCT GTA GAA ATT AAT TGT ACA AGA Thr Ile Ile Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg 80 85 90	288
55	CCC AAC AAC AAT ACA AGA AAA AGA ATC CGT ATC CAG AGA GGA CCA GGG Pro Asn Asn Asn Thr Arg Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly 95 100 105 110	336
55	AGA GCA TTT GTT ACA ATA GGA AAA ATA GGA AAT ATG AGA CAA GCA CAT Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His 115 120 125	384

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

5 Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro
1 5 10 15
Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
20 25 30
10 Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg
35 40 45
Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
50 55 60
15 Glu Val Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile
65 70 75 80
Ile Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
80 85 90 95
Asn Asn Thr Arg Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala
100 105 110
25 Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn
115 120 125
Ile Ser Arg Ala Lys Trp Asn Asn Thr Leu Lys Gln Ile Asp Ser Lys
130 135 140
30 Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser
145 150 155 160
Ser Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly
165 170 175
35 Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe
180 185 190
40 Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp
195 200 205
Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln
210 215 220
45 Lys Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg
225 230 235 240
Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn
245 250 255
50 Ser Asn Asn Glu Ser Glu Ile Phe Arg Leu Gly Gly Gly Asp Met Arg
260 265 270
55 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
275 280 285

Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg
290 295 300

5 Glu Lys Arg Ala Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly
305 310 315 320

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln
325 330 335

10 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu
340 345 350

15 Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly
355 360 365

Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys
370 375 380

20 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys
385 390 395 400

Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu
405 410 415

25 Gln Ile Trp Asn His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn
420 425 430

30 Asn Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln
435 440 445

Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser
450 455 460

35 Leu Trp Asp Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu
465 470 475 480

Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala
485 490 495

40 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser
500 505 510

Phe Gln Thr His Leu Pro Ile

45 515

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CACCCCTCTC CTACGTAACC AAGGATC

27

5 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTACTGGTCA CCATATTGGT CAAC

24

20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAGAGAGAT GGGAGCTCGA GCGTC

25

35

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

40 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCCCCCTAT ACGTATTGTG

20

50

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

55 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCAGTGAATT CCTAATACGA CTCACTATAG GTTAAAACAG C

41

10

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCTATCCTG AGCTCCATAT GTGTCGAGCA GTTTTTGGTT TAGCATTG

48

25

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: peptide

35

(v) FRAGMENT TYPE: internal

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Thr Lys Asp Leu Thr Thr Tyr Gly
1 5

45

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2220 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: cDNA

55

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..2203

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

5	CGA CCA GCA GAC CAG ACA GTC ACA GCA GCC TTG ACA AAA CGT TCC TGG	48
	Arg Pro Ala Asp Gln Thr Val Thr Ala Ala Leu Thr Lys Arg Ser Trp	
	1 5 10 15	
10	AAC TCA AGC ACT TCT CCA CAG AGG AGG ACA GAG CAG ACA GCA GAG ACC	96
	Asn Ser Ser Thr Ser Pro Gln Arg Arg Thr Glu Gln Thr Ala Glu Thr	
	20 25 30	
15	ATG GAG TCT CCC TCG GCC CCT CCC CAC AGA TGG TGC ATC CCC TGG CAG	144
	Met Glu Ser Pro Ser Ala Pro Pro His Arg Trp Cys Ile Pro Trp Gln	
	35 40 45	
20	AGG CTC CTG CTC ACA GCC TCA CTT CTA ACC TTC TGG AAC CCG CCC ACC	192
	Arg Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr	
	50 55 60	
25	ACT GCC AAG CTC ACT ATT GAA TCC ACG CCG TTC AAT GTC GCA GAG GGG	240
	Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly	
	65 70 75 80	
30	AAG GAG GTG CTT CTA CTT GTC CAC AAT CTG CCC CAG CAT CTT TTT GGC	288
	Lys Glu Val Leu Leu Leu Val His Asn Leu Pro Gln His Leu Phe Gly	
	85 90 95	
35	TAC AGC TGG TAC AAA GGT GAA AGA GTG GAT GGC AAC CGT CAA ATT ATA	336
	Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Ile	
	100 105 110	
40	GGA TAT GTA ATA GGA ACT CAA CAA GCT ACC CCA GGG CCC GCA TAC AGT	384
	Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser	
	115 120 125	
45	GGT CGA GAG ATA ATA TAC CCC AAT GCA TCC CTG CTG ATC CAG AAC ATC	432
	Gly Arg Glu Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Ile	
	130 135 140	
50	ATC CAG AAT GAC ACA GGA TTC TAC ACC CTA CAC GTC ATA AAG TCA GAT	480
	Ile Gln Asn Asp Thr Gly Phe Tyr Thr Leu His Val Ile Lys Ser Asp	
	145 150 155 160	
55	CTT GTG AAT GAA GAA GCA ACT GGC CAG TTC CGG GTA TAC CCG GAG CTG	528
	Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu	
	165 170 175	
60	CCC AAG CCC TCC ATC TCC AGC AAC AAC TCC AAA CCC GTG GAG GAC AAG	576
	Pro Lys Pro Ser Ile Ser Ser Asn Asn Ser Lys Pro Val Glu Asp Lys	
	180 185 190	
65	GAT GCT GTG GCC TTC ACC TGT GAA CCT GAG ACT CAG GAC GCA ACC TAC	624
	Asp Ala Val Ala Phe Thr Cys Glu Pro Glu Thr Gln Asp Ala Thr Tyr	
	195 200 205	
70	CTG TGG TGG GTA AAC AAT CAG AGC CTC CCG GTC AGT CCC AGG CTG CAG	672
	Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg Leu Gln	

TTT ATC GCC AAA ATC ACG CCA AAT AAT AAC GGG ACC TAT GCC TGT TTT 2064
Phe Ile Ala Lys Ile Thr Pro Asn Asn Asn Gly Thr Tyr Ala Cys Phe
675 680 685

5 GTC TCT AAC TTG GCT ACT GGC CGC AAT AAT TCC ATA GTC AAG AGC ATC 2112
Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Ser Ile
690 695 700

10 ACA GTC TCT GCA TCT GGA ACT TCT CCT GGT CTC TCA GCT GGG GCC ACT 2160
Thr Val Ser Ala Ser Gly Thr Ser Pro Gly Leu Ser Ala Gly Ala Thr
705 710 715 720

15 GTC GGC ATC ATG ATT GGA GTG CTG GTT GGG GTT GCT CTG ATA 2202
Val Gly Ile Met Ile Gly Val Leu Val Gly Val Ala Leu Ile
725 730

TAGCAGCCCT GGTGTAGT 2220

20 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 734 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Arg Pro Ala Asp Gln Thr Val Thr Ala Ala Leu Thr Lys Arg Ser Trp
1 5 10 15

Asn Ser Ser Thr Ser Pro Gln Arg Arg Thr Glu Gln Thr Ala Glu Thr
20 25 30

Met Glu Ser Pro Ser Ala Pro Pro His Arg Trp Cys Ile Pro Trp Gln
35 40 45

40 Arg Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
50 55 60

Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly
65 70 75 80

Lys Glu Val Leu Leu Leu Val His Asn Leu Pro Gln His Leu Phe Gly
85 90 95

50 Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Ile
100 105 110

Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser
115 120 125

55 Gly Arg Glu Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Ile
130 135 140

	Ile	Gln	Asn	Asp	Thr	Gly	Phe	Tyr	Thr	Leu	His	Val	Ile	Lys	Ser	Asp	
	145					150					155						160
5	Leu	Val	Asn	Glu	Glu	Ala	Thr	Gly	Gln	Phe	Arg	Val	Tyr	Pro	Glu	Leu	
				165						170					175		
	Pro	Lys	Pro	Ser	Ile	Ser	Ser	Asn	Asn	Ser	Lys	Pro	Val	Glu	Asp	Lys	
				180					185					190			
10	Asp	Ala	Val	Ala	Phe	Thr	Cys	Glu	Pro	Glu	Thr	Gln	Asp	Ala	Thr	Tyr	
			195					200					205				
	Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	Leu	Gln	
15		210					215					220					
	Leu	Ser	Asn	Gly	Asn	Arg	Thr	Leu	Thr	Leu	Phe	Asn	Val	Thr	Arg	Asn	
		225				230					235					240	
	Asp	Thr	Ala	Ser	Tyr	Lys	Cys	Glu	Thr	Gln	Asn	Pro	Val	Ser	Ala	Arg	
20					245					250					255		
	Arg	Ser	Asp	Ser	Val	Ile	Leu	Asn	Val	Leu	Tyr	Gly	Pro	Asp	Ala	Pro	
				260					265					270			
25	Thr	Ile	Ser	Pro	Leu	Asn	Thr	Ser	Tyr	Arg	Ser	Gly	Glu	Asn	Leu	Asn	
			275					280					285				
	Leu	Ser	Cys	His	Ala	Ala	Ser	Asn	Pro	Pro	Ala	Gln	Tyr	Ser	Trp	Phe	
30			290					295				300					
	Val	Asn	Gly	Thr	Phe	Gln	Gln	Ser	Thr	Gln	Glu	Leu	Phe	Ile	Pro	Asn	
		305				310					315					320	
	Ile	Thr	Val	Asn	Asn	Ser	Gly	Ser	Tyr	Thr	Cys	Gln	Ala	His	Asn	Ser	
35					325					330					335		
	Asp	Thr	Gly	Leu	Asn	Arg	Thr	Thr	Val	Thr	Thr	Ile	Thr	Val	Tyr	Ala	
				340					345					350			
40	Glu	Pro	Pro	Lys	Pro	Phe	Ile	Thr	Ser	Asn	Asn	Ser	Asn	Pro	Val	Glu	
			355					360					365				
	Asp	Glu	Asp	Ala	Val	Ala	Leu	Thr	Cys	Glu	Pro	Glu	Ile	Gln	Asn	Thr	
45			370				375					380					
	Thr	Tyr	Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	
		385				390					395					400	
	Leu	Gln	Leu	Ser	Asn	Asp	Asn	Arg	Thr	Leu	Thr	Leu	Leu	Ser	Val	Thr	
50					405					410					415		
	Arg	Asn	Asp	Val	Gly	Pro	Tyr	Glu	Cys	Gly	Ile	Gln	Asn	Glu	Leu	Ser	
				420					425					430			
55	Val	Asp	His	Ser	Asp	Pro	Val	Ile	Leu	Asn	Val	Leu	Tyr	Gly	Pro	Asp	
			435					440					445				

0076664 0000000

```

Asp Pro Thr Ile Ser Pro Ser Tyr Thr Tyr Tyr Arg Pro Gly Val Asn
450                               455                               460

5  Leu Ser Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser
465                               470                               475                               480

Trp Leu Ile Asp Gly Asn Ile Gln Gln His Thr Gln Glu Leu Phe Ile
                               485                               490                               495

10 Ser Asn Ile Thr Glu Lys Asn Ser Gly Leu Tyr Thr Cys Gln Ala Asn
                               500                               505                               510

Asn Ser Ala Ser Gly His Ser Arg Thr Thr Val Lys Thr Ile Thr Val
                               515                               520                               525

15 Ser Ala Glu Leu Pro Lys Pro Ser Ile Ser Ser Asn Asn Ser Lys Pro
                               530                               535                               540

Val Glu Asp Lys Asp Ala Val Ala Phe Thr Cys Glu Pro Glu Ala Gln
20 545                               550                               555                               560

Asn Thr Thr Tyr Leu Trp Trp Val Asn Gly Gln Ser Leu Pro Val Ser
                               565                               570                               575

25 Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu Phe Asn
                               580                               585                               590

Val Thr Arg Asn Asp Ala Arg Ala Tyr Val Cys Gly Ile Gln Asn Ser
                               595                               600                               605

30 Val Ser Ala Asn Arg Ser Asp Pro Val Thr Leu Asp Val Leu Tyr Gly
                               610                               615                               620

Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Ser Ser Tyr Leu Ser Gly
35 625                               630                               635                               640

Ala Asn Leu Asn Leu Ser Cys His Ser Ala Ser Asn Pro Ser Pro Gln
                               645                               650                               655

40 Tyr Ser Trp Arg Ile Asn Gly Ile Pro Gln Gln His Thr Gln Val Leu
                               660                               665                               670

Phe Ile Ala Lys Ile Thr Pro Asn Asn Gly Thr Tyr Ala Cys Phe
45 675                               680                               685

Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Ser Ile
                               690                               695                               700

Thr Val Ser Ala Ser Gly Thr Ser Pro Gly Leu Ser Ala Gly Ala Thr
50 705                               710                               715                               720

Val Gly Ile Met Ile Gly Val Leu Val Gly Val Ala Leu Ile
                               725                               730

```

55 (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

10 CCAGTGAATT CCTAATACGA CTACCTATAG GTTAAACAG C

41

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

25 GATGAACCCT CGAGACCCAT TATG

24

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: cDNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCACCAAGTA CGTAACCACA TATGG

25

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTGAGGACTG CTGG

14

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

15 CACCACTGCC CTCGAGAAGC TCACTATTG

29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

30 CACCACTGCC CTCGAGAAGC TCACTATTG

29